

RAW SEQUENCE LISTING

**The Biotechnology Systems Branch of the Scientific and Technical
Information Center (STIC) no errors detected.**

Application Serial Number: 10/686,835 B
Source: IFW16
Date Processed by STIC: 10/18/2005

ENTERED



IFW16

RAW SEQUENCE LISTING

DATE: 10/18/2005

PATENT APPLICATION: US/10/686,835B

TIME: 08:55:41

Input Set : A:\sequence listing.txt

Output Set: N:\CRF4\10182005\J686835B.raw

4 <110> APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.
6 <120> TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM
7 HEPATITIS C VIRUS
9 <130> FILE REFERENCE: 13/083-2-C1
11 <140> CURRENT APPLICATION NUMBER: 10/686,835B
12 <141> CURRENT FILING DATE: 2003-10-16
14 <150> PRIOR APPLICATION NUMBER: 10/029,907
15 <151> PRIOR FILING DATE: 2001-12-21
17 <150> PRIOR APPLICATION NUMBER: 60/257,857
18 <151> PRIOR FILING DATE: 2000-12-22
20 <160> NUMBER OF SEQ ID NOS: 28
22 <170> SOFTWARE: FastSEQ for Windows Version 4.0
24 <210> SEQ ID NO: 1
25 <211> LENGTH: 8639
26 <212> TYPE: DNA
27 <213> ORGANISM: HCV
29 <220> FEATURE:
30 <221> NAME/KEY: CDS
31 <222> LOCATION: (1803)...(8408)
33 <400> SEQUENCE: 1
34 ggccagcccc cgattggggg cgacactcca ccatagatca ctcccctgtg aggaactact 60
35 gtcttcacgc agaaagcgtc tagccatggc gttagtatga gtgtcgtgca gcctccagga 120
36 cccccctcc cgaggagagcc atagtgggtct gcggaaccgg tgagtacacc ggaattgcca 180
37 ggacgaccgg gtcctttctt ggatcaaccc gctcaatgcc tggagatttg ggcgtgcccc 240
38 cgcgagactg ctagccgagt agtggtgggt cgcgaaaggg cttgtggtac tgctgatag 300
39 ggtgcttgcg agtgccccgg gaggtctcgt agaccgtgca ccatgagcac gaatcctaaa 360
40 cctcaaagaa aaaccaaagg gcgcgccatg attgaacaag atggattgca cgcaggttct 420
41 ccggccgctt ggggtggagag gctattcggc tatgactggg cacaacagac aatcggctgc 480
42 tctgatgccg ccgtgttccg gctgtcagcg caggggcgcc cggttctttt tgtcaagacc 540
43 gacctgtccg gtgcctgaa tgaactgcag gacgaggcag cgcggtatc gtggctggcc 600
44 acgacgggcg ttccttgccg agctgtgctc gacgttgctc ctgaagcggg aagggactgg 660
45 ctgctatttg gcgaagtgcc ggggcaggat ctctgtcat ctcacctgc tcctgccgag 720
46 aaagtatcca tcatggctga tgcaatgcgg cggctgcata cgcttgatcc ggctacctgc 780
47 ccattcgacc accaagcgaa acatcgcac gagcgagcac gtactcggat ggaagccggg 840
48 cttgtcgatc aggatgatct ggacgaagag catcaggggc tcgcgccagc cgaactgttc 900
49 gccaggctca aggcgcgcac gcccgacggc gaggatctcg tcgtgaccca tggcgatgcc 960
50 tgcttgccga atatcatggg ggaatatggc cgcttttctg gattcatcga ctgtggccgg 1020
51 ctgggtgtgg cggaccgcta tcaggacata gcgttggtc cccgtgatat tgctgaagag 1080
52 cttggcgggc aatgggctga ccgcttctc gtgctttacg gtatcgccgc tcccgattcg 1140
53 cagcgcacgc cttctatcg cttcttgac gagttcttct gagttcgcgc ccagatgtta 1200
54 acagaccaca acggtttccc tctagcggga tcaattccgc cccccccct aacgttactg 1260
55 gccgaagccg cttggaataa ggccggtgtg cgtttgtcta tatgttattt tccaccatat 1320
56 tgccgtcttt tggcaatgtg agggcccgga aacctggccc tgtcttcttg acgagcattc 1380

Cpg-6)

RAW SEQUENCE LISTING

DATE: 10/18/2005

PATENT APPLICATION: US/10/686,835B

TIME: 08:55:41

Input Set : A:\sequence listing.txt

Output Set: N:\CRF4\10182005\J686835B.raw

```

57 ctaggggtct tccccctctc gccaaaggaa tgcaaggtct gttgaatgtc gtgaaggaaag 1440
58 cagttcctct ggaagcttct tgaagacaaa caacgtctgt agcgaccctt tgcaggcagc 1500
59 ggaaccccccc acctggcgac aggtgcctct gcggccaaaa gccacgtgta taagatacac 1560
60 ctgcaaaggc ggcacaaccc cagtgccacg ttgtgagttg gatagttgtg gaaagagtca 1620
61 aatggctctc ctcaagcgta ttcaacaagg ggctgaagga tgcccagaag gtacccatt 1680
62 gtatgggatac tgatctgggg cctcggtgca catgctttac atgtgtttag tgcagggttaa 1740
63 aaaacgtcta ggccccccga accacgggga cgtggttttc ctttgaaaaa cacgataata 1800
64 cc atg gac cgg gag atg gca gca tgc tgc gga ggc gcg gtt ttc gta 1847
65 Met Asp Arg Glu Met Ala Ala Ser Cys Gly Gly Ala Val Phe Val
66 1 5 10 15
68 ggt ctg ata ctc ttg acc ttg tca ccg cac tat aag ctg ttc ctc gct 1895
69 Gly Leu Ile Leu Leu Thr Leu Ser Pro His Tyr Lys Leu Phe Leu Ala
70 20 25 30
72 agg ctc ata tgg tgg tta caa tat ttt atc acc agg gcc gag gca cac 1943
73 Arg Leu Ile Trp Trp Leu Gln Tyr Phe Ile Thr Arg Ala Glu Ala His
74 35 40 45
76 ttg caa gtg tgg atc ccc ccc ctc aac gtt cgg ggg ggc cgc gat gcc 1991
77 Leu Gln Val Trp Ile Pro Pro Leu Asn Val Arg Gly Gly Arg Asp Ala
78 50 55 60
80 gtc atc ctc ctc acg tgc gcg atc cac cca gag cta atc ttt acc atc 2039
81 Val Ile Leu Leu Thr Cys Ala Ile His Pro Glu Leu Ile Phe Thr Ile
82 65 70 75
84 acc aaa atc ttg ctc gcc ata ctc ggt cca ctc atg gtg ctc cag gct 2087
85 Thr Lys Ile Leu Leu Ala Ile Leu Gly Pro Leu Met Val Leu Gln Ala
86 80 85 90 95
88 ggt ata acc aaa gtg ccg tac ttc gtg cgc gca cac ggg ctc att cgt 2135
89 Gly Ile Thr Lys Val Pro Tyr Phe Val Arg Ala His Gly Leu Ile Arg
90 100 105 110
92 gca tgc atg ctg gtg cgg aag gtt gct ggg ggt cat tat gtc caa atg 2183
93 Ala Cys Met Leu Val Arg Lys Val Ala Gly Gly His Tyr Val Gln Met
94 115 120 125
96 gct ctc atg aag ttg gcc gca ctg aca ggt acg tac gtt tat gac cat 2231
97 Ala Leu Met Lys Leu Ala Ala Leu Thr Gly Thr Tyr Val Tyr Asp His
98 130 135 140
100 ctc acc cca ctg cgg gac tgg gcc cac gcg ggc cta cga gac ctt gcg 2279
101 Leu Thr Pro Leu Arg Asp Trp Ala His Ala Gly Leu Arg Asp Leu Ala
102 145 150 155
104 gtg gca gtt gag ccc gtc gtc ttc tct gat atg gag acc aag gtt atc 2327
105 Val Ala Val Glu Pro Val Val Phe Ser Asp Met Glu Thr Lys Val Ile
106 160 165 170 175
108 acc tgg ggg gca gac acc gcg gcg tgt ggg gac atc atc ttg ggc ctg 2375
109 Thr Trp Gly Ala Asp Thr Ala Ala Cys Gly Asp Ile Ile Leu Gly Leu
110 180 185 190
112 ccc gtc tcc gcc cgc agg ggg agg gag ata cat ctg gga ccg gca gac 2423
113 Pro Val Ser Ala Arg Arg Gly Arg Glu Ile His Leu Gly Pro Ala Asp
114 195 200 205
116 agc ctt gaa ggg cag ggg tgg cga ctc ctc gcg cct att acg gcc tac 2471
117 Ser Leu Glu Gly Gln Gly Trp Arg Leu Leu Ala Pro Ile Thr Ala Tyr
118 210 215 220

```

RAW SEQUENCE LISTING

DATE: 10/18/2005

PATENT APPLICATION: US/10/686,835B

TIME: 08:55:41

Input Set : A:\sequence listing.txt

Output Set: N:\CRF4\10182005\J686835B.raw

```

120 tcc caa cag acg cga ggc cta ctt ggc tgc atc atc act agc ctc aca 2519
121 Ser Gln Gln Thr Arg Gly Leu Leu Gly Cys Ile Ile Thr Ser Leu Thr
122      225      230      235
124 ggc cgg gac agg aac cag gtc gag ggg gag gtc caa gtg gtc tcc acc 2567
125 Gly Arg Asp Arg Asn Gln Val Glu Gly Glu Val Gln Val Val Ser Thr
126 240      245      250      255
128 gca aca caa tct ttc ctg gcg acc tgc gtc aat ggc gtg tgt tgg act 2615
129 Ala Thr Gln Ser Phe Leu Ala Thr Cys Val Asn Gly Val Cys Trp Thr
130      260      265      270
132 gtc tat cat ggt gcc ggc tca aag acc ctt gcc ggc cca aag ggc cca 2663
133 Val Tyr His Gly Ala Gly Ser Lys Thr Leu Ala Gly Pro Lys Gly Pro
134      275      280      285
136 atc acc caa atg tac acc aat gtg gac cag gac ctc gtc ggc tgg caa 2711
137 Ile Thr Gln Met Tyr Thr Asn Val Asp Gln Asp Leu Val Gly Trp Gln
138      290      295      300
140 gcg ccc ccc ggg gcg cgt tcc ttg aca cca tgc acc tgc ggc agc tcg 2759
141 Ala Pro Pro Gly Ala Arg Ser Leu Thr Pro Cys Thr Cys Gly Ser Ser
142      305      310      315
144 gac ctt tac ttg gtc acg agg cat gcc gat gtc att ccg gtg cgc cgg 2807
145 Asp Leu Tyr Leu Val Thr Arg His Ala Asp Val Ile Pro Val Arg Arg
146 320      325      330      335
148 cgg ggc gac agc agg ggg agc cta ctc tcc ccc agg ccc gtc tcc tac 2855
149 Arg Gly Asp Ser Arg Gly Ser Leu Leu Ser Pro Arg Pro Val Ser Tyr
150      340      345      350
152 ttg aag ggc tct tcg ggc ggt cca ctg ctc tgc ccc tcg ggg cac gct 2903
153 Leu Lys Gly Ser Ser Gly Gly Pro Leu Leu Cys Pro Ser Gly His Ala
154      355      360      365
156 gtg ggc atc ttt cgg gct gcc gtg tgc acc cga ggg gtt gcg aag gcg 2951
157 Val Gly Ile Phe Arg Ala Ala Val Cys Thr Arg Gly Val Ala Lys Ala
158      370      375      380
160 gtg gac ttt gta ccc gtc gag tct atg gaa acc act atg cgg tcc ccg 2999
161 Val Asp Phe Val Pro Val Glu Ser Met Glu Thr Thr Met Arg Ser Pro
162      385      390      395
164 gtc ttc acg gac aac tcg tcc cct ccg gcc gta ccg cag aca ttc cag 3047
165 Val Phe Thr Asp Asn Ser Ser Pro Pro Ala Val Pro Gln Thr Phe Gln
166 400      405      410      415
168 gtg gcc cat cta cac gcc cct act ggt agc ggc aag agc act aag gtg 3095
169 Val Ala His Leu His Ala Pro Thr Gly Ser Gly Lys Ser Thr Lys Val
170      420      425      430
172 ccg gct gcg tat gca gcc caa ggg tat aag gtg ctt gtc ctg aac ccg 3143
173 Pro Ala Ala Tyr Ala Ala Gln Gly Tyr Lys Val Leu Val Leu Asn Pro
174      435      440      445
176 tcc gtc gcc gcc acc cta ggt ttc ggg gcg tat atg tct aag gca cat 3191
177 Ser Val Ala Ala Thr Leu Gly Phe Gly Ala Tyr Met Ser Lys Ala His
178      450      455      460
180 ggt atc gac cct aac atc aga acc ggg gta agg acc atc acc acg ggt 3239
181 Gly Ile Asp Pro Asn Ile Arg Thr Gly Val Arg Thr Ile Thr Thr Gly
182      465      470      475
184 gcc ccc atc acg tac tcc acc tat ggc aag ttt ctt gcc gac ggt ggt 3287

```

RAW SEQUENCE LISTING

DATE: 10/18/2005

PATENT APPLICATION: US/10/686,835B

TIME: 08:55:41

Input Set : A:\sequence listing.txt

Output Set: N:\CRF4\10182005\J686835B.raw

```

185 Ala Pro Ile Thr Tyr Ser Thr Tyr Gly Lys Phe Leu Ala Asp Gly Gly
186 480          485          490          495
188 tgc tct ggg ggc gcc tat gac atc ata ata tgt gat gag tgc cac tca 3335
189 Cys Ser Gly Gly Ala Tyr Asp Ile Ile Ile Cys Asp Glu Cys His Ser
190          500          505          510
192 act gac tcg acc act atc ctg ggc atc ggc aca gtc ctg gac caa gcg 3383
193 Thr Asp Ser Thr Thr Ile Leu Gly Ile Gly Thr Val Leu Asp Gln Ala
194          515          520          525
196 gag acg gct gga gcg cga ctc gtc gtg ctc gcc acc gct acg cct ccg 3431
197 Glu Thr Ala Gly Ala Arg Leu Val Val Leu Ala Thr Ala Thr Pro Pro
198          530          535          540
200 gga tcg gtc acc gtg cca cat cca aac atc gag gag gtg gct ctg tcc 3479
201 Gly Ser Val Thr Val Pro His Pro Asn Ile Glu Glu Val Ala Leu Ser
202          545          550          555
204 agc act gga gaa atc ccc ttt tat ggc aaa gcc atc ccc atc gag acc 3527
205 Ser Thr Gly Gly Ile Pro Phe Tyr Gly Lys Ala Ile Pro Ile Glu Thr
206 560          565          570          575
208 atc aag ggg ggg agg cac ctc att ttc tgc cat tcc aag aag aaa tgt 3575
209 Ile Lys Gly Gly Arg His Leu Ile Phe Cys His Ser Lys Lys Lys Cys
210          580          585          590
212 gat gag ctc gcc gcg aag ctg tcc ggc ctc gga ctc aat gct gta gca 3623
213 Asp Glu Leu Ala Ala Lys Leu Ser Gly Leu Gly Leu Asn Ala Val Ala
214          595          600          605
216 tat tac cgg ggc ctt gat gta tcc gtc ata cca act agc gga gac gtc 3671
217 Tyr Tyr Arg Gly Leu Asp Val Ser Val Ile Pro Thr Ser Gly Asp Val
218          610          615          620
220 att gtc gta gca acg gac gct cta atg acg ggc ttt acc ggc gat ttc 3719
221 Ile Val Val Ala Thr Asp Ala Leu Met Thr Gly Phe Thr Gly Asp Phe
222          625          630          635
224 gac tca gtg atc gac tgc aat aca tgt gtc acc cag aca gtc gac ttc 3767
225 Asp Ser Val Ile Asp Cys Asn Thr Cys Val Thr Gln Thr Val Asp Phe
226 640          645          650          655
228 agc ctg gac ccg acc ttc acc att gag acg acg acc gtg cca caa gac 3815
229 Ser Leu Asp Pro Thr Phe Thr Ile Glu Thr Thr Val Pro Gln Asp
230          660          665          670
232 gcg gtg tca cgc tcg cag cgg cga ggc agg act ggt agg ggc agg atg 3863
233 Ala Val Ser Arg Ser Gln Arg Arg Gly Arg Thr Gly Arg Gly Arg Met
234          675          680          685
236 ggc att tac agg ttt gtg act cca gga gaa cgg ccc tcg ggc atg ttc 3911
237 Gly Ile Tyr Arg Phe Val Thr Pro Gly Glu Arg Pro Ser Gly Met Phe
238          690          695          700
240 gat tcc tcg gtt ctg tgc gag tgc tat gac gcg ggc tgt gct tgg tac 3959
241 Asp Ser Ser Val Leu Cys Glu Cys Tyr Asp Ala Gly Cys Ala Trp Tyr
242          705          710          715
244 gag ctc acg ccc gcc gag acc tca gtt agg ttg cgg gct tac cta aac 4007
245 Glu Leu Thr Pro Ala Glu Thr Ser Val Arg Leu Arg Ala Tyr Leu Asn
246 720          725          730          735
248 aca cca ggg ttg ccc gtc tgc cag gac cat ctg gag ttc tgg gag agc 4055
249 Thr Pro Gly Leu Pro Val Cys Gln Asp His Leu Glu Phe Trp Glu Ser

```

RAW SEQUENCE LISTING

DATE: 10/18/2005

PATENT APPLICATION: US/10/686,835B

TIME: 08:55:41

Input Set : A:\sequence listing.txt

Output Set: N:\CRF4\10182005\J686835B.raw

250		740		745		750	
252	gtc ttt aca ggc ctc acc cac ata gac gcc cat ttc ttg tcc cag act						4103
253	Val Phe Thr Gly Leu Thr His Ile Asp Ala His Phe Leu Ser Gln Thr						
254		755		760		765	
256	aag cag gca gga gac aac ttc ccc tac ctg gta gca tac cag gct acg						4151
257	Lys Gln Ala Gly Asp Asn Phe Pro Tyr Leu Val Ala Tyr Gln Ala Thr						
258		770		775		780	
260	gtg tgc gcc agg gct cag gct cca cct cca tcg tgg gac caa atg tgg						4199
261	Val Cys Ala Arg Ala Gln Ala Pro Pro Pro Ser Trp Asp Gln Met Trp						
262		785		790		795	
264	aag tgt ctc ata cgg cta aag cct acg ctg cac ggg cca acg ccc ctg						4247
265	Lys Cys Leu Ile Arg Leu Lys Pro Thr Leu His Gly Pro Thr Pro Leu						
266	800		805		810		815
268	ctg tat agg ctg gga gcc gtt caa aac gag gtt act acc aca cac ccc						4295
269	Leu Tyr Arg Leu Gly Ala Val Gln Asn Glu Val Thr Thr Thr His Pro						
270		820		825		830	
272	ata acc aaa tac atc atg gca tgc atg tcg gct gac ctg gag gtc gtc						4343
273	Ile Thr Lys Tyr Ile Met Ala Cys Met Ser Ala Asp Leu Glu Val Val						
274		835		840		845	
276	acg agc acc tgg gtg ctg gta ggc gga gtc cta gca gct ctg gcc gcg						4391
277	Thr Ser Thr Trp Val Leu Val Gly Gly Val Leu Ala Ala Leu Ala Ala						
278		850		855		860	
280	tat tgc ctg aca aca ggc agc gtg gtc att gtg ggc agg atc atc ttg						4439
281	Tyr Cys Leu Thr Thr Gly Ser Val Val Ile Val Gly Arg Ile Ile Leu						
282		865		870		875	
284	tcc gga aag ccg gcc atc att ccc gac agg gaa gtc ctt tac ccg gag						4487
285	Ser Gly Lys Pro Ala Ile Ile Pro Asp Arg Glu Val Leu Tyr Arg Glu						
286	880		885		890		895
288	ttc gat gag atg gaa gag tgc gcc tca cac ctc cct tac atc gaa cag						4535
289	Phe Asp Glu Met Glu Glu Cys Ala Ser His Leu Pro Tyr Ile Glu Gln						
290		900		905		910	
292	gga atg cag ctc gcc gaa caa ttc aaa cag aag gca atc ggg ttg ctg						4583
293	Gly Met Gln Leu Ala Glu Gln Phe Lys Gln Lys Ala Ile Gly Leu Leu						
294		915		920		925	
296	caa aca gcc acc aag caa gcg gag gct gct gct ccc gtg gtg gaa tcc						4631
297	Gln Thr Ala Thr Lys Gln Ala Glu Ala Ala Ala Pro Val Val Glu Ser						
298		930		935		940	
300	aag tgg ccg acc ctc gaa gcc ttc tgg gcg aag cat atg tgg aat ttc						4679
301	Lys Trp Arg Thr Leu Glu Ala Phe Trp Ala Lys His Met Trp Asn Phe						
302		945		950		955	
304	atc agc ggg ata caa tat tta gca ggc ttg tcc act ctg cct ggc aac						4727
305	Ile Ser Gly Ile Gln Tyr Leu Ala Gly Leu Ser Thr Leu Pro Gly Asn						
306	960		965		970		975
308	ccc gcg ata gca tca ctg atg gca ttc aca gcc tct atc acc agc ccg						4775
309	Pro Ala Ile Ala Ser Leu Met Ala Phe Thr Ala Ser Ile Thr Ser Pro						
310		980		985		990	
312	ctc acc acc caa cat acc ctc ctg ttt aac atc ctg ggg gga tgg gtg						4823
313	Leu Thr Thr Gln His Thr Leu Leu Phe Asn Ile Leu Gly Gly Trp Val						
314		995		1000		1005	

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/10/686,835B

DATE: 10/18/2005
TIME: 08:55:42

Input Set : A:\sequence listing.txt
Output Set: N:\CRF4\10182005\J686835B.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:2; Xaa Pos. 882,1489

Seq#:3; Xaa Pos. 882,1489

VERIFICATION SUMMARY

DATE: 10/18/2005

PATENT APPLICATION: US/10/686,835B

TIME: 08:55:42

Input Set : A:\sequence listing.txt

Output Set: N:\CRF4\10182005\J686835B.raw

L:634 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:2
L:638 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:2
L:890 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2 after pos.:4489
M:341 Repeated in SeqNo=2
L:1240 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:3
L:1351 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3 after pos.:880
M:341 Repeated in SeqNo=3